

E-Discovery, Analytics, and The Business of Legal Operations

Jeanne Somma - Director of Legal Innovation &
Discovery Counsel, RVM Enterprises



HOFSTRA UNIVERSITY®

MAURICE A. DEANE SCHOOL OF LAW

Agenda

- **Introduction – Why Do We Need to Understand E-Discovery?**
- **E-Discovery Reference Model (EDRM)**
- **Overview of Analytics and Predictive Coding (TAR)**
- **Corporate Legal Operations Overview**
- **Merging Legal Practice, Technology, and Business Acumen**

Why do we need to understand E-Discovery?

“[We are currently in]...a time when the amount of available data for discovery is growing exponentially. More data has been created in the last two years than in the entire previous history of the human race and the amount of data is predicted to grow 10-fold by 2020.”

Gordon v. T.G.R. Logistics, Inc., No. 16-CV-00238-NDF, 2017 WL 1947537 (D. Wyo. May 10, 2017).

E-Discovery IS Discovery

- The largest percentage of Discovery centers around Electronically Stored Information (ESI)
- Attorneys must understand what documents a party possesses and may use to support its claims or defenses as well as all aspects of those documents
 - This includes metadata (data about data), retention rules of that data, special treatment the data may require, and other technical aspects.
- Initial disclosures must be made within 14 days after the 26(f) conference so speed to insight is key.
- As document volumes balloon lawyers must adapt strategies to quickly identify and review pertinent information using Search Terms or TAR for example.
- Savvy plaintiffs, counsel and various government agencies are extremely aggressive in taking control of discussions and making unreasonable demands regarding ESI.

How Much Data Do I Have

- CD = 650 MB = ~50,000 pages
- DVD = 4.7 GB = ~350,000 pages
- 1 MB is about ~75 pages
- 1 GB is about ~75,000 pages
- Average pages per email: 1.5 (100,099 pages per GB)
- Average pages per Word document: 8 (64,782 pages per GB)
- Average pages per Excel spreadsheet: 50 (165,791 pages per GB)
- Average pages per PowerPoint: 14 (17,552 pages per GB)
- 100 MB PST file is 900 emails and 300 attachments
- 600 MB PST file is 5,500 emails and 1,600 attachments
- A 1.0 GB NSF file is 9,000 emails and 3,000 attachments



HOFSTRA UNIVERSITY®

MAURICE A. DEANE SCHOOL OF LAW

E-Discovery, Technology, and Competence

- **Lawyers have a Duty of Technological Competence**

- ABA Model Rules 1.1: Competence

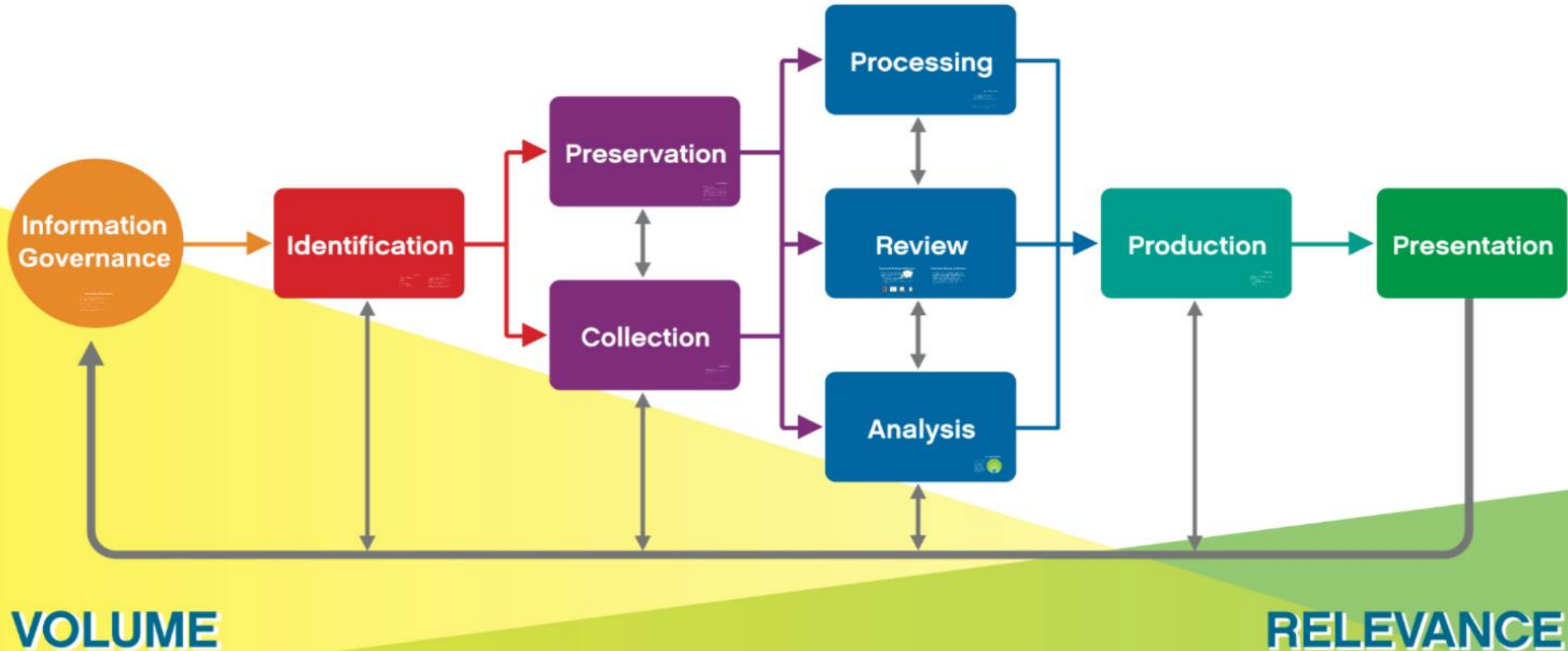
- ABA's Model Rules were modified in 2012 to confirm that a lawyer's duty of competence requires keeping "abreast of changes in the law and its practice," which includes knowing "the benefits and risks and associated with relevant technology."
 - Lawyers are expected to have at least a basic understanding of the technologies they use, the risks associated with those technologies and the means available to mitigate those risks.

- Currently over half of all jurisdictions have enacted rules mandating that attorneys become and remain familiar with technologies that may impact their practices.

E-Discovery and Updated Federal Rules

- FRCP 16: Courts expect you to be ready for litigation, including being fluent in the IT and network architecture, so that the pretrial conference leads to agreements on what ESI is discoverable. FRCP 26(f) sanctions for not obeying a scheduling or pretrial order are a good thing to avoid.
- FRCP 26: Provides protection from excessive or expensive e-discovery requests, except when you don't deserve that protection.
 - FRCP 26(a)(1)(C): Requires that you make initial disclosures no later than 14 days after the Rule 26(f) meet and confer, unless an objection or another time is set by stipulation or court order. If you have an objection, now is the time to voice it.
 - Rule 26(b)(2)(B): Introduced the concept of not reasonably accessible ESI. The concept of not reasonably accessible paper had not existed. This rule provides procedures for shifting the cost of accessing not reasonably accessible ESI to the requesting party.
 - FRCP 26(b)(5)(B): Gives courts a clear procedure for settling claims when you hand over ESI to the requesting party that you shouldn't have.
 - Rule 26(f): This is the meet and confer rule. This rule requires all parties to meet within 99 days of the lawsuit's filing and at least 21 days before a scheduled conference.
 - Rule 26(g): Requires an attorney to sign every e-discovery request, response, or objection.
- FRCP 34: Establishes a structured way to resolve disputes over document production.
 - FRCP 34(b): Establishes protocols for how documents are produced to requesting parties. As the requesting party, you choose the form of production. Most often, the requested form is native file because those files tend to reveal the most. You might not have the equipment or expertise to read the produced ESI easily if it's not in native form or a form you pick. This is usually a matter of negotiation between the parties.
- FRCP 37: Judges have the power, courtesy of Rule 37(f), to impose sanctions against a party “who fails to obey an order to provide or permit discovery.”
 - Rule 37(e): Creates a safe harbor from sanctions if you did not preserve, and therefore no longer have, ESI that's requested provided that certain conditions and circumstances are met. Judges also have powers that are considered inherent in the court that expand the ability to impose sanctions beyond Rule 37.
- FRCP 45: If you're a nonparty to e-discovery, you're protected from some of the costs or burdens that parties typically have to pay or endure.
- FRE 502: Protects attorney-client privilege and provides some protection against inadvertent disclosure, if you're quick enough to notice your mistake and meet other conditions.
- FRE 502(b): If attorney-client privileged or work product protected material is inadvertently disclosed, you might be able to get it back if you took reasonable steps to prevent the error; and noticed and responded promptly to fix the error.
- FRE 901: Requires that ESI, like physical evidence, be authenticated to verify that it is what it claims to be. Metadata may be used to authenticate an ESI.

The E-Discovery Reference Model



Overview of Automation, Analytics and Predictive Coding (TAR)

Analytics Defined

- the science of logical analysis
- the analysis of data, typically large sets of business data, by the use of mathematics, statistics, and computer software
- the patterns and other meaningful information gathered from the analysis of data

Analytics – Types and Uses

Structured Analytics (Content)

- Email Threading
- Duplicate Detection
- Language Detection
- Geographic Analysis
- Timeline Analysis
- Communication & Domain Analysis

Unstructured Analytics (Context)

- Conceptual
 - Keyword Expansion
 - Clustering
- Textual Classification
 - Categorization
 - Predictive Coding/TAR
 - Junk Exclusion
- Tone/Emotion (Concern)



HOFSTRA UNIVERSITY®

MAURICE A. DEANE SCHOOL OF LAW

Predictive Coding (TAR)

Predictive Coding is a process of having computer software electronically classify documents based on input from expert reviewers, in an effort to expedite the organization and prioritization of the document collection.

The computer classification may include broad topics pertaining to discovery responsiveness, privilege, and other designated issues.

Predictive Coding may dramatically reduce the time and cost of reviewing ESI, by reducing the amount of human review needed on documents classified as potentially non-material.

Key Concepts

- Unsupervised and Supervised Learning
- Richness
- Precision
- Recall
- F Measurement
- Simple Passive Learning
- Simple Active Learning
- Continuous Ranking
- Continuous Active Learning



HOFSTRA UNIVERSITY®

MAURICE A. DEANE SCHOOL OF LAW

TAR 1.0 and TAR 2.0

TAR 1.0

- Set Goals
- Code the Control Set
- Receive and Review Training Sets Starting with the Seed
- Predict Results
- Test Results
- Finalize the Process
- Review the Identified Population

Limitations

Lacks Feedback Opportunities

- Once the team gets going on the review set, there is no opportunity to feed back their judgments on review documents and improve the ranking algorithm.

Requires Subject Matter Experts (SME)

- Generally requires a senior lawyer or SME for training. Expert training can require the review of thousands of documents to build a control set, train and then test the results.

Introduction of New Documents

- New documents render the control set invalid because they were not part of the random selection process.

Low richness collections

- It can be hard to find good training examples based on random sampling. If richness is below 1 percent, several thousand additional documents may need to be reviewed just to find enough relevant ones to train the system.

TAR 2.0

- Sample and code documents.
 - These documents are for the initial training of the algorithm.
 - Documents may be found through initial searches, through witness interviews or perhaps from earlier reviews.
- Review Documents
 - As the review progresses, coded documents are continuously fed to the algorithm to continue the training.
- Continue the review process
 - To the point that the number of relevant documents decreases substantially or runs out.
- Sample the un-reviewed documents
 - Use this sample to estimate the percentage of relevant documents found and those remaining. If the recall percentage is sufficient for your purposes, you can stop the review.



HOFSTRA UNIVERSITY®

MAURICE A. DEANE SCHOOL OF LAW

Corporate Legal Operations Overview

Legal Operations is a multi-disciplinary function that optimizes legal services delivery to a business or government entity by focusing on twelve core competencies.



HOFSTRA UNIVERSITY

MAURICE A. DEANE SCHOOL OF LAW

CLOC Survey

Quick Stats

50%

Of law dept's have a preferred provider program

3

median number of years law dept's have had a preferred provider program

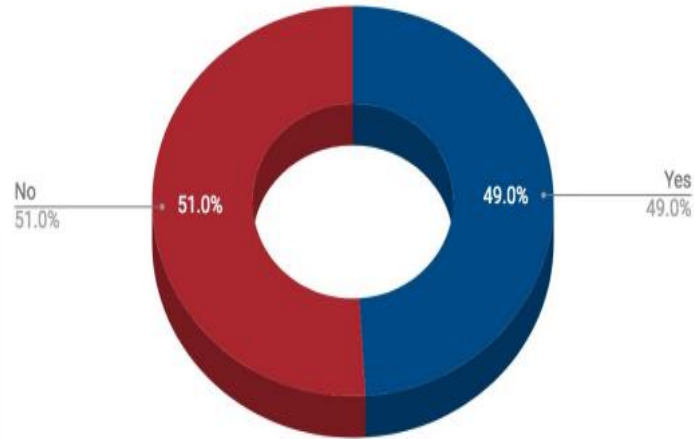
12%

Of law dept's work exclusively with firms in their preferred provider program

23%

average estimated savings for law dept's with a preferred provider program

Do you have a preferred provider program?



Top three reasons for creating preferred provider program:

1) To secure better rates

2) To save money

3) To secure better spend predictability

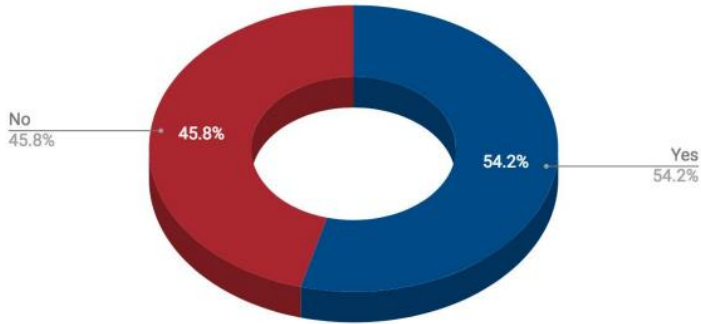
OVERALL, RESPONDENTS LEVERAGE **74%** OF THEIR OUTSIDE SPEND WITH PREFERRED FIRMS

38% of respondents have saved **20%+** through their preferred provider program.

CLOC Survey

In House Firm Selection Requirements

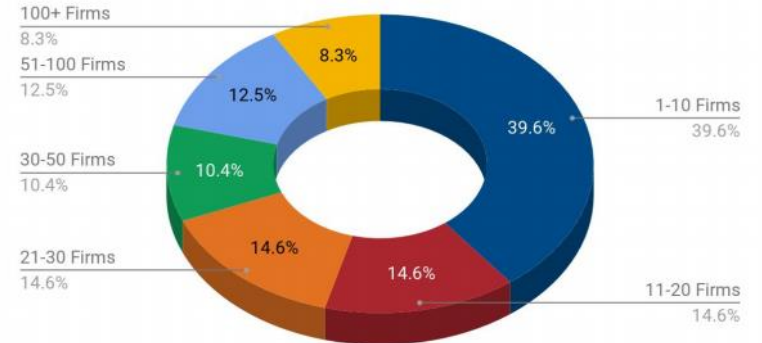
Do you have a policy defining when your preferred panel firms must be used?



13
Average number of substantive areas of law covered by preferred provider firms

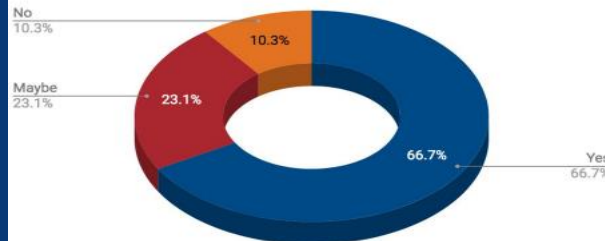
Number of Firms in Program

How many firms are in your program?

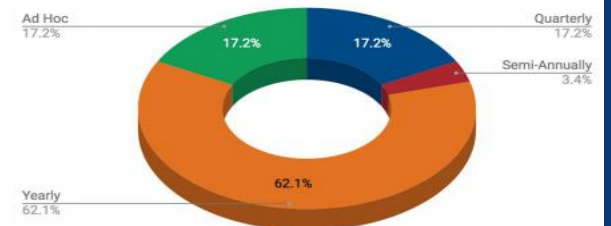


Business Reviews of Preferred Provider Program Firms

Do you perform business reviews with firms in your program?



How often do you review firms in your program?



Pulling it All Together

We succeed where we add value.

Behave Like a Lawyer, Think Like a Businessperson

The value of our services lies in the ability to actively listen to a problem and find a beneficial solution that limits exposure for our client. To do that we must be able to focus our energy on strategic thinking and action.

On the other hand there will always be administrative and lower level tasks that demand our attention. We must strive to automate those as best as possible in order to focus our efforts on the tasks that add value to our client and keep us billable.

Lastly we must constantly assess our performance, but personally and company-wide. Where there are no analytics and no metrics performance cannot be managed. If you cannot manage your performance then you cannot use key performance indicators to help improve and grow.